REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
09/02/2012	Final	Jan., 1, 2009-Nov., 30, 2011

4. TITLE AND SUBTITLE
Scalable effective approaches for quadratic assignment problems based on conic optimization and applications

5a. CONTRACT NUMBER FA9550-09-1-0098

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jiming Peng and Hans Mittlemann

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Industrial and Enterprise System Engineering
University of Illinois at Urbana-Champaign
104 S. Mathews Ave.
Urbana, IL 61801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Mathematics, Information, and Life Sciences
Air Force Office of Scientific Research
875 N. Randolph St,
Arlington, VA 22203-1768

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0352

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approval for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the authors and should not be construed as an official U.S. Government position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This project deals with quadratic assignment problems (QAP) that arise from a broad range of applications such as target tracking, resource allocation and communications. The main endeavor of the project is to explore the structure of the associated data matrix of the underlying problem, combining with various matrix splitting schemes, to derive strong and yet cheap to compute convex optimization relaxations. The new relaxation can be further used to obtain a good approximation to the original problem or help the development of exact algorithms. The efficacy of the proposed approaches has been demonstrated via both theoretical and numerical comparisons with other existing approaches in the literature, and the techniques developed through the project have been successfully applied to large scale QAPs from communications and information processing.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
u	u	u	u	1	Jiming Peng
					19b. TELEPHONE NUMBER *(Include area code)* 217-2446275

Reset

Third Annual and Final Report on AFOSR grant FA9550-09-1-0098

Scalable effective approaches for quadratic assignment problems based on conic optimization and applications

----

PI: Jiming Peng University of Illinois at Urbana-Champaign

Co-PI: Hans Mittelmann Arizona State University

Project Summary

During the period (2009-2011) of the above grant, the PI and Co-PI have worked on the following issues pertaining to the project:

1) Introducing a new framework to construct semidefinite optimization (SDO) relaxations for quadratic assignment problems (QAP) based on various matrix splitting schemes. More specifically, we introduced numerous approaches to explore the structure in the associated data matrix of the underlying QAP to obtain a strong lower bound and compare the new relaxations with other existing relaxation models for QAPs in the literature from both a theoretical and computational perspectives. Our research in this direction has led to the following publications:

   [1] J. Peng, H. Mittleman and X. Li. A new convex relaxation framework for quadratic assignment problems based on matrix splitting. Math. Programming: Computation, 1(2), 59-77, 2010.

   [2] H. Mittleman and J. Peng. Estimating bounds for quadratic assignment problems associated with Hamming and Manhattan distance matrices based on semidefinite programming. SIAM J. Optim. Volume 20(6), pp. 3408-3426, 2010.

   [3] J. Peng, T. Zhu, H.Zh. Luo and K.Ch. Toh. Semi-definite Relaxation of Quadratic Assignment Problems based on Non-redundant Matrix Splitting. Under review for SIAM J. Optim. Submitted in 2011.

2) Developing effective approximation algorithms for numerous notorious hard and fundamental discrete optimization problems that arise from a broad range of applications such as the densest k-subgraph problems and the binary matrix factorization. Publiscations in this direction include the following papers:

[4] M. Heath, P. Jiang, J. Peng and R. Yang. Finding the densest k-subgraph via 1-mean clustering and low-dimension approximation. Under review for Mathematical programming: Computation. Submitted in 2010.

[5] P. Jiang, J. Peng and M. Heath. Alternating Update Procedures for Unconstrained and Constrained Binary Matrix Factorization. Under review for the 18^{th} ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2012).

3) Applying the new techniques developed in this project to solve or find a good approximation to large scale problems from communications and information theory. Publications in this direction include the following papers:

[6] X. Wu, H.D. Mittelmann, X. Wang, and J. Wang, On Computation of Performance Bounds of Optimal Index Assignment, Data Compression Conference (DCC) 2010. IEEE, 189-198 (2010), DOI 10.1109/DCC.2010.24.

[7] X. Wu, H.D. Mittelmann, X. Wang, and J. Wang, On Computation of Performance Bounds of Optimal Index Assignment, IEEE Trans Comm 59(12), 3229-3233 (2011) . DOI: 10.1109/TCOMM.2011.081111.100300.

[8] D. C. Gijswijt, H. D. Mittelmann, and A. Schrijver.Semidefinite code bounds based on quadruple distances. To appear in IEEE Transactions on Information Theory, DOI 10.1109/TIT.2012.2184845.

4) In addition to the above three research subjects outlined in the initial proposal, we have also worked on numerous miscellaneous topics that were not listed in the original proposal yet related to the project. These include the development of several optimization models and algorithm design for problems from computer vision and learning, research on sparse solutions in quadratic optimization with random data, and SDP relaxations for several hard discrete optimization problems. Publications on these miscellaneous topics include the following papers:

[9] L. Mukherjee, V. Singh, J. Peng and C. Hinrichs. Learning kernels for variants of normalized cuts: Convex relaxations and applications, Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR), June 2010. (acceptance 27%).

[10] L. Mukherjee, V. Singh, J. Peng. Scale invariant cosegmentation for image groups. Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR), June 2011. (Oral presentation, acceptance ratio 3.5%).

[11] X. Chen, J. Peng and Sh. Zhang. Sparse solutions to random standard quadratic optimization problems. To appear in Mathematical Programming, 2012.

[12] H. D. Mittelmann and F. Vallentin, High Accuracy Semidefinite Programming Bounds for Kissing Numbers, Exper. Math. 19, 174-179 (2010).

One PhD student, Mr. Tao Zhu in UI, has been supported by this grant. Tao had successfully passed the comprehensive exam in the ISE department of UI in the year of 2010. Tao has already finished one research paper on how to find better matrix splitting schemes to improve the SDP relaxation model for QAPs. The paper (see reference [3]) has been submitted to SIAM J. Optimization for possible publication. He has also made substantial progress in developing SDO-based approaches for robust optimization. Tao is expected to graduate in the year of 2013.

To summarize, during the project period, the PI and co-PI have worked together on most of the major issues listed in the original proposal and successfully realized most research objectives of the project. Three papers have been published or accepted by leading journals in optimization such as mathematical programming and SIAM J. Optimization, and 2 other papers are under review for MP and SIOPT. We also developed effective approximation algorithms for numerous notorious hard discrete optimization problems and the corresponding papers have been submitted for publications in the top conferences or journals in their disciplines. We have also applied the new techniques developed in this project to problems from communications and information thory. The corresponding papers [7,8] have been accepted by top journals in these two disciplines. The publication of the results of the project in top journals in optimization and other related disciplines demonstrates the success of the project.

The project has also met its educational goal. The PhD student supported by the project has made substantial progress in his study and research and will graduate as originally scheduled.

One issue listed in the original proposal but we did not provide a satisfactory answer is the development of a branch-bound (B-B) approach for large-scale QAPs. The PI and Co-PI, however, did implement a prototype version of the B-B approach (using the new SDO relaxation) and compared with several existing B-B approaches in the literature. We also observed significant improvement in solving the relaxed sub-problem in every node of the B-B approach. Nevertheless, since the B-B approach usually takes a tremendous time to solve instances of size n=30, while the major purpose of this project is to obtain strong bounds or good approximation to large scale QAPs of size a few hundreds or even thousands, therefore, we decided to concentrate on the development of scalable approaches for large scale problem.

More details on PI and Co-PI's research in the year of 2011 will be documented in the following pages separately.

Third Annual Report on AFOSR grant FA9550-09-1-0098

PI: Jiming Peng University of Illinois at Urbana-Champaign

During the third period of the above grant, the PI has working on several issues listed in the project summary. In what follows we describe more details regarding the progress we made in these topics respectively.

1: New SDO relaxations for QAPs based on non-redundant matrix splitting

Our work in this direction is some extension of the results in papers [1,2]. Note that though the new SDP relaxation models for QAPs introduced in [1,2] have a much smaller size compared with existing SDP relaxations for the same problem in the literature, it still remains open how to a better matrix splitting scheme such that the resulting SDO relaxation is the strongest among various splitting schemes. For this, we consider the splitting B=B1-B2, where both B1 and B2 are positive semi-definite. We then introduced a new notion called non-redundant splitting, which is a necessary condition for the optimal splitting whose resulting bound is the strongest. We then propose to follow the so-called minimal trace principle to find such a non-redundant splitting and showed that some splitting schemes in [1] and [2] are non-redundant splitting. New non-redundant matrix splitting schemes are proposed and experiment results show that the new SDO relaxation from non-redundant matrix splitting can indeed help to improve the bound for most QAP instances in the literature. For more details, see paper [3].

2. Effective approximation to unconstrained and constrained binary matrix factorization (UBMF/CBMF)

UBMF refers to the problem of finding two low-rank matrices such that the difference between a given binary matrix and the matrix product of two-low-rank matrices is minimal to some extent.
Such a problem arises frequently in data analysis of data sets with binary attributes where a low-dimensional representation is crucial to find meaning patterns. The model was first introduced in the following paper:

> [Ref1] M. Koyuturk and A. Grama. PROXIMUS: a framework for analyzing very high dimensionaldiscrete-attributed datasets. In ACM SIGKDD, 147-156, 2003.

Since then, numerous approaches have been proposed for the UBMF problem. Nevertheless, due to the hardness of the problem, approximation algorithm has only been reported for the special case where the matrix is of rank 1. See

[Ref2] B. H. Shen, S. Ji, and J. Ye. Mining discrete patterns via binary matrix factorization. In ACM SIGKDD, 757-766, 2009.

It should be pointed out that the algorithm in the above paper uses the lifting-and-projection to the rank-1 matrix product to reformulate it as a large scale binary linear programming (LP), and then finds a 2-approximation via solving its LP relaxation. The expensive lifting-projection procedure and the rank-1 restriction prevent it from many applications.

To cope with theU BMF problem, in [5], we first considered a more restricted version, the CBMF problem and explored the interrelation between CBMF and UBMF. Then, we introduced two specific variants of CBMF and established the equivalence between the variants of CBMF and data clustering. Effective approximation algorithms were proposed for the CBMF. As a byproduct, we also obtained effective approximation to the UBMF problem. For example, for the rank-1 case, the proposed algorithms runs in $O(n^2+m^2)$ time where n and m is the size of the two rank-1 matrices. As illustrated by experimental results in [5], the new algorithm significantly over performed other existing algorithms in the literature for the BMF problem in term of both the quality of the obtained solution and the CPU time to obtain such a solution. For more details, see [5].

3: Sparse solutions to random standard quadratic programming problems (StQPs)

StQP refers to the problem of minimizing a quadratic form over the simplex and such a problem arises from numerous domains such as control, resource allocation and social network. The StQP provides a prototype for many classes of QPs and the problem has been proved to be NP-hard. However, it has been observed in some recent works [9] that global optimal solutions to StQPs from various applications can be located easily and the solution is usually to be very sparse. In [11], we first explored the optimality conditions of the StQP to derive an intrinsic link between the global solution to an StQP and some probability events. Then, we used order statistics to estimate the probability that a random StQP has an optimal solution with at least k nonzero elements. This gives a precise probabilistic characterization of the sparsity of the sparsest optimal solution to StQPs. The results in [11] provide new theoretical insights on StQPs and their extensions and help to understand some issues in QPs that have been observed for long. Those theoretical insights can also help the design of algorithm by exploring the sparsity in the solution.

Third and Final Annual Report on AFOSR grant FA9550-09-1-0098

PI Hans Mittelmann, Arizona State University

During the third period of the above grant our research had concentrated on the following three issues:

1) Solving very large 2d-QAPs in a real-life problem from communications

It is well known that important QAPs arise in communications. In fact, the only known real-life instance of the multidimensional QAP is found in this area, namely in the implementation of a hybrid ARQ (automatic repeat request) scheme for enriching diversity among multiple packet re-transmissions, by optimizing the mapping of data bits to modulation symbols. The three-dimensional QAP had been considered in

[Ref3] P.Hahn et al, The quadratic three-dimensional assignment problem: Exact and approximate solution methods, Europ. J. of Oper. Res., 184, 416-428 (2008)

We have chosen a real-life 2-d QAP from communications where sizes are $n = 2^d$, $d = 2,..,9$. The larger sizes, n=128, 256, 512 were either impossible for us before or were requiring excessive resources, or the bounds produced were not strong. Together with researchers from communications we addressed the problem of optimal index assignment in VQ (vector quantization) relevant to minimizing channel distortion in wireless communication, a topic of great practical importance.

A first refereed conference proceedings paper was [6]. Two major improvements had been made to the results in [6] during the third grant period (see [7]).

i) The size n=512, a very formidable size for computing quality lower or upper bounds, has been solved.

ii) In [6] only the case of the adjacency matrix of the hypercube was treated. This amounts to the communication having only 1-bit errors. We have generalized to multiple bit errors.

The following two tables are from [7].

Here GLB denotes the Gilmore-Lawler lower bound, proj. bd. the well-known, also among communication engineers, projection bound. The upper boumd was computed by us utilizing an approach introduced by T. Stuetzle. While the gaps between our lower bound and this upper bound does not seem to be so small, they amount to less than 1.78 dB in the first and less than 1.82 dB in the second table. These are very small gaps compared to state-of-the-art knowledge in comunications.

	32	64	128	256	512
GLB	84784	58617	45503	43942	38156
proj bd	102156	42807	neg.	neg.	neg.
our l.b.	304551	289883	242657	199959	193271
upper b.	358984	349337	334360	294756	291314

Table 1. Bounds for adjacency matrix of the hypercube (1-bit errors)

	32	64	128	256	512
GLB	884	819	484	465	416
proj bd	1146	614	neg.	neg.	neg.
our l.b.	2891	2834	2412	2020	1931
upper b.	3608	3734	3347	2984	2937

Table 2. Bounds for general case (multiple bit errors)

It must be noted that, as described in [7] the straightforward application of our bounding methods developed in the joint work with PI Peng did not produce good bounds and we found a variant that did.

Topic 2) Solving the related problem of binary codes of maximal Hamming distance

In the above communication problem as is customary in communication engineering a code book is chosen, that is a subset of all binary words of length n and then codewords are assigned based on maximizing Hamming distance. Fundamental is thus how many binary words of length

n there are that have a minimal Hamming distance, of say d bits.
This is one of the open questions in combinatorics and apart from
some small dimensional cases, these numbers are unknown. An up-to-date
table of the best known bounds is kept at

[Ref3] Table of general binary codes http://www.win.tue.nl/~aeb/codes/binary-1.html

In collaboration with Lex Schrijver and Dion Gijswijt we have greatly
improved the known upper bounds. Again, difficult to solve but powerful
semidefinite relaxations played a key role. The paper [8] has been accepted.
There were several improved bounds added to the table during the third year
of the grant and before publication. In this paper a substantial portion of the known
bounds were improved although not all of them as we were able to do in the related paper
[12].

Topic 3) Continuing work on computing bounds for higher dimensional QAPs

Several researchers have made progress on solving exactly small
three-dimensional QAP problems, see, for example [Ref3]. Mainly a group in France had
applied both bounding techniques and B&B. However, the methods employed will not be
applicable to larger cases. Their lower bounding technique is a variant of Hahn's RLT-1
method. Solving, for example, the nug16a/b instances from QAPLIB took 28/31 days.

The most important instance of the higher dimensional QAP is the axial QAP. It has
complexity (n!)^(m-1) in dimension m. The repeated transmission problem in
communications mentioned above is a typical example. In each transmission one adds
another n! factor from the underlying 2d QAP. So there are at least two sources for higher
dimensional QAPs. Generation through repeating well-known 2d QAPs such as from
QAPLIB and from the real-life communication example in our research.

During the last year of the grant we made substantial progress on solving
QAPs of any dimension. The method will be described in forthcoming papers. Suffice it
to say that we were able to leverage all we had done for the
2d QAP in a way that works well for QAPLIB derived instances and which
in future research we plan to extend to the communication problem.

In assessing the quality of our lower bounds for m-diemnsional QAPs it is
crucial to also have a strong upper bounding technique available. As in
our work [7,8] we successfully applied the iterated local search (ILS)
method of Stuetzle also used in [Ref3].

In order to have a benchmark problem we contacted Bertrand LeCun who in

their poject CHOC from 2005-2008 had applied their parallel B&B framework BOB++ to the RLT1 technique of Peter Hahn and had obtained exact solutions up to dimension 16, namely for the two QAPLIB-generated axial Q3APs NUG16a and NUG16b. They reported optimal values of 15132 for NUG16a and 10498 for NUG16b. They also sent us one set of optimal permutations for NUG16a.

Applying our method we obtained bounds of 11202 and 7300 respectively.
Peter Hahn applied his RLT1 method and reported to us bounds of 2318 and 803 after 8.5 hours CPU time. The French group may have modified the RLT1 method and have started with a better bound.

Here are sample CPU times.
lobd is our lower bound and exact are the times from Bertrand LeCun.

	lobd	exact	Hahn lobd
NUG16a	10 secds	30.6 days	8.5 hrs
NUG16b	12 secds	27 days	8.5 hrs
NUG30	260 secds		
SKO42	45 min		

Table 3. CPU times for our lower bounds and times needed by others

All nugent problems from QAPLIB were then extended to three dimensions. This table lists the lower and upper bounds obtained.

NUG	lobd	upbd
12	2776	4174
14	5587	8084
15	6686	9336
16a	11202	15132
16b	7300	10498
17	11767	16142
18	13020	17964
20	18712	25590
21	15604	17210
22	29113	38503
24	26536	29632
25	28812	32085
27	45562	61570
28	42682	53381
30	53235	67973

Table 4. Our lower bounds and the ILS upper bounds for Q3APs from NUG problems

We used the SKO problems from QAPLIB for somewhat larger sizes.

SKO	lobd	upbd
42	183143	230888
64	799201	953994
81	1828730	2089170

Table 5. Our lower bounds and the ITS upper bounds for Q3APs from SKO problems

Since this test was so successful we used QAPLIB generated higher dimensional QAPs to test our method. We have obtained lower bounds in dimension m for problems of size n:

m	n
3	81
4	64
5	49
6	42
7	36

Table 6. Sizes of SKO problems solved by us in dimension m

While the corresponding bounds for the communication problem still have to be obtained, the above results clearly show our ability to solve higher dimensional QAPs.

All our papers cited above are accessible through
http://plato.asu.edu/papers.html